# A generalist vision–language foundation model for diverse biomedical tasks

## Abstract

Traditional biomedical artificial intelligence (AI) models, designed for specific tasks or modalities, often exhibit limited flexibility in real-world deployment and struggle to utilize holistic information. Generalist AI holds the potential to address these limitations due to its versatility in interpreting different data types and generating tailored outputs for diverse needs. However, existing biomedical generalist AI solutions are typically heavyweight and closed source to researchers, practitioners and patients. Here, we describe BiomedGPT, the first open-source and lightweight vision–language foundation model, designed as a generalist capable of performing various biomedical tasks. BiomedGPT achieved state-of-the-art results in 16 out of 25 experiments while maintaining a computing-friendly model scale. We also conducted human evaluations to assess the capabilities of BiomedGPT in radiology visual question answering, report generation and summarization. BiomedGPT exhibits robust prediction ability with a low error rate of 3.8% in question answering, satisfactory performance with an error rate of 8.3% in writing complex radiology reports, and competitive summarization ability with a nearly equivalent preference score to human experts. Our method demonstrates that effective training with diverse data can lead to more practical biomedical AI for improving diagnosis and workflow efficiency. An open-source and computing-friendly vision–language model achieves state-of-the-art accuracy in 16 out of 25 biomedical tasks, with promising performance in a series of potential clinical applications.